Know Thy Toucher

Dominik Schmidt
Computing Department, Lancaster University
Lancaster, UK
schmidtd@comp.lancs.ac.uk

ABSTRACT

Most of current academic and commercial surface computing systems are capable of multitouch detection and hence allow simultaneous input from multiple users. Although there are so far only few applications in this area which rely on identifying the user, we believe that the association of touches to users will become an essential feature of surface computing as applications mature, new application areas emerge, and the enabling technology is readily available. As the capacitive technology used in present user identification enabled tabletops is limited with respect to the supported number of users and screen size, we outline a user identification enabled tabletop concept based on computer vision and biometric hand shape information, and introduce the prototype system we built to further investigate this concept. In a preliminary consideration, we derive concepts for identifying users by examining what new possibilities are enabled and by introducing different scopes of identification.

Author Keywords

Multitouch, multiuser, surface computing, tabletop, user identification.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Besides numerous advances in fundamental system technology, researchers have been exploring new interaction techniques and applications especially tailored to interactive multitouch surfaces. Without doubt, the fascination of surface computing stems in parts from its directness of interaction and of course its multitouch and hence multiuser capabilities. Instead of having only the single point of control which a traditional personal computer offers with its mouse and keyboard interface, users can now interact simultaneously and collaborate equitably.

On the application side, this new accumulation of parallel input information poses new challenges, ranging from understanding spatially and temporally related touches and interpreting gestures to differentiating users in order to enable more advanced interaction techniques. However, although much of the tabletop research is focused on multiuser interactions [11], surprisingly few publications deal with identifying the user and using information beyond the actual points of touch.

We believe that identifying the user who is interacting is certainly a huge benefit and will become an essential information as applications mature, new application areas emerge, and the enabling technology is readily available. In conjunction with user-aware applications, user identification enables powerful new interaction techniques. Some of the applications applying DiamondTouch technology [3, 12, 15, 16] show on a small scale what is possible if touches can be associated to users. Furthermore, previous work in the area of single- or shared-display groupware (SDG) [1, 8, 13] has explored diverse user-aware concepts and implementations illustrating the benefits of user identification. In addition, there already exists some work on more general user-aware applications and identification concepts [14].

A common scenario that illustrates the importance of user identification on the basis of independent modes of interaction is a collaborative drawing application with a shared whiteboard [1, 8, 11]: Two or more users are sitting around a tabletop to work together on a drawing, only using their fingers to interact with the application. From the available tool palette, they can chose different line widths and colors. As the users work simultaneously on different parts of the drawing, they use different line widths and colors at the same time. That is, if one of the users selected red he or she expects the next stroke to be red while the other users might have selected blue. In order to realize this behavior, the application needs to know who selected a mode and who is initiating an action, i.e. touches have to be associated to users.

In contrast to previous SDG systems, direct touch surface systems do not need to employ electronic devices for interaction. While not having such a mediator makes the interaction more instantaneous and natural, the user identification becomes unlike more challenging because there is no active instance which could transmit an identifier. Currently, there seems to exist only one available tabletop system which supports user identification, namely DiamondTouch [3, 6].

As the capacitive technology used in DiamondTouch is rather limited in terms of supported number of users, screen size, and mobility of users around the table (users have to be in contact with a conductive plate in order to make the system recognize touches), a more flexible surface computing solution with user identification capabilities is desirable to ease the development of new interaction techniques.

Several current systems use a computer vision approach to successfully detect touches. Given that computer vision facilitates scaling to larger screen sizes and is contactless, an investigation is needed to explore how this technology can be applied to identify users. With the goal to enable user identification for computer vision-based tabletop systems, we derive concepts for identifying users by examining what new possibilities are enabled and by introducing different levels of identification as a guideline for developing a system in a bottom-up approach, i.e. to realize the least complicated identification tasks first. Moreover, we provide an outline for a user identification enabled tabletop system based on hand contour extraction and biometric concepts. To conduct further research in this area, we built a rear-projected tabletop system using direct illumination which we equipped with an overhead camera for hand tracking.

RELATED WORK

Stewart et. al. [13] define „Single Display Groupware (SDG) to be computer programs that enable co-present users to collaborate via a shared computer with a single shared display and simultaneous use of multiple input devices." From the context, it can be assumed that every input device is capable of representing one cursor only. Therefore, actions can be associated to a single user. A collaborative drawing application is described where only the user who selected a tool to work with (e.g. a brush) is allowed to change its properties such as color or line width.

The „multi-device multi-user multi-editor" [1] was designed to be used by groups sharing a single screen; users can edit text and draw rectangles. The system supports multiple pointing devices that are registered to a user in order to enable user identification. Registration is done by taking one of the input devices and clicking on a name bar. Once registered, the user's preferences are available. Different users can select different modes and work simultaneously within the same work space, even on the same object. Users are represented by colors which are used for displaying cursor, insertion points, and selections.

Pebbles [8] is a project which investigates how portable computers can be used to interact with PCs. Parts of the research deals with SDG where multiple users, each of them equipped with a handheld device, work together on a shared display controlling cursor and keyboard. Either a single cursor is used and users have to take turns or a separate cursor exists for every user. As every handheld device runs its own client application and communicates with a server connected to the shared display, cursors can be associated with devices. PebblesDraw, a shared whiteboard application, allows users to draw simultaneously in different modes. It was designed with the goal that „each user's actions are independent". New users are registered by specifying the communication port for their handheld device; hence their actions can easily be associated later. PebblesDraw supports a separate cursor per user. Widgets have been adopted for simultaneous multiuser usage: When a mode is selected, it is not shown in the toolbar (which would be confusing as the toolbar is shared by everyone) but in the user's home area as well as next to the respective cursor. Undo-by-user is supported in addition to global undo.

Tse et. al. [14] developed a toolkit to facilitate the management of input from multiple mice and keyboards for rapidly prototyping SDG. SDGToolkit supports rotation of the devices' input space to accommodate for different orientations in tabletop setups. Associating cursors to users is straight forward as every event is associated to an input device and hence accompanied by an identifier. An interface for developing user-aware controls is provided.

DiamondTouch [3] is a tabletop platform which supports user identification. It was developed with the goal to enable „simultaneous and spontaneous" interaction rather than interacting in a turn-taking style. One of the six characteristics the authors list is „identifying", i.e. the table „detects which user is touching each point". In order to demonstrate the technology, a multiplayer game was implemented. The players have to pop bubbles of their color by touching them to score points; points are deducting for destroying the other players' bubbles. Besides this game which makes use of the ability to associate touches to users the authors point out the concepts of „virtual personal work areas" and „privileged objects" which only respond to the corresponding user. The DiamondTouch technology supports up to four users [6].

Some of the applications developed for DiamondTouch make use of identifying the user, such as „RoomPlaner" [16] – a two-user room furniture layout application which divides the application space into one public and two private areas. Leveraging the table's user identification capabilities, private areas can only be manipulated by their respective owner. UbiTable [12] is another tabletop application using DiamondTouch which allows users to bring documents from their personal devices to the shared table. As the table defines areas with different privacy policies and as documents have an owner associated to certain rights it is crucial that the user who is interacting is identified in order to decide whether a certain operation is allowed or not. Tse et. al. [15] investigate how commercial single user applications can be adopted for multiuser tabletops. Gesture input is combined with speech input to realize a multimodal interface. Here, user identification is needed to correctly associate speech and gesture commands.

iDwidgets [9] is a conceptual framework consisting of a set of reusable GUI elements which rely on user identification to provide customization in four dimensions: function, content, appearance, and group input. A button can for example execute different commands for different users. In addition, knowing who is touching enables modal input sequences and the interleaving of different people's activities. Taking the abstract concepts to practice, several example implementations of the concepts discussed before are presented [10].

Carreira et. al. [2] extract the hands' rotations from a camera mounted over a tabletop as complement to fingertip touches. Touches are not associated to users; however, the hand's orientation is used to enable new interaction techniques. Dohse et. al. [4] augment a rear projection touch table with an overhead camera to track hands in order to, first, assign touches to users and, second, increase robustness by removing noise.

CONCEPTS FOR USER IDENTIFICATION

In the following, we use the concept of user identification in a broader sense, i.e. we do not necessarily aim at revealing the user's true identity. However, we are certainly interested in assigning touches to users and distinguishing different users which already enables new possibilities for interaction techniques as discussed below. If the user is additionally known to the system, the mapping of input activity to identities enables role-based concepts.

New Possibilities

As can be seen from the examples provided in the beginning, several applications already apply user identification. In this context, it seems to be worthwhile exploring which possibilities are enabled by associating touches to users with a focus on surface computing systems.

Filtering On a very basic level, we can handle unrelated touches appropriately by only considering those touches which are assigned to a user. Other touches might have been caused by noise inherent to the respective technology, by other objects placed on the table surface, or by the user's arm lying on the surface during an interaction.

Advanced Gestures Gestures involving multiple fingers or hands can be interpreted unambiguously to enable more advanced interaction techniques. Moreover, gestures involving more than one user, such as hand-over gestures, can be realized.

User-Aware Interaction User-aware interfaces enable modal input sequences as illustrated in the scenario described in the introduction. Moreover, further concepts identified in the iD-Widgets framework [9] under the label „customizing group input“ can be implemented, such as a voting system where a widget might require a certain amount of users to activate it.

Roles If the interface is not only aware of different users but can also associate them to known roles, „customizing function“, „customizing content“, and „customizing appearance“ concepts as identified in the context of the iDwidgets framework [9] can be realized.

Hence, it is possible to grant access to widgets for certain users only, realizing private areas for example. Moreover, different functions can be executed depending on the initiating user or the displayed information is personalized for the user sending a query.

Scope of Identification

In order to identify the requirements for implementing these new possibilities for a vision-based system, we suggest a bottom-up approach. On the basis of key questions relating to the scope of identification, requirements are discussed with increasing complexity in terms of the underlying computer vision tasks.

Does a touch originate from a user's hand? To answer this question, hand contours and touches need to be identified in the input image and set into relation: If a touch lies within the shape of a hand it originated from this hand. Answering this question enables us to perform *filtering* operations as described in the previous section.

Do two simultaneous touches originate from the same hand? The answer to this question links otherwise unrelated touches. Having this information at hand, we cannot say that two simultaneous touches are *not* from the same user – as he or she might use both hands. However, we know for sure that two simultaneous touches are from the same user if they are from the same hand. Applying the approach discussed in the previous question, it is straight forward to decide if two simultaneous touches originate from the same hand by differentiating between hand shapes. This enables the use of advanced gestures in parts – at least as long as these gestures only involve multiple fingers and not multiple hands or users.

Do any two touches originate from the same user? In order to enable *advanced gestures* involving multiple hands or users and realize *user-aware interactions*, hand shapes for different users need to be distinguished or clustered. Users may for example wear an accessory, such as a finger ring or a wrist watch, equipped with an optical transmitter that sends out coded light pulses to be picked up by the installed camera. However, to realize a less obtrusive system, we suggest to extract features out of the hand contours in order to be able to differentiate between hands of different users. Moreover, by keeping a database which maps hand features to users *roles* can be implemented.

Work-in-Progress: A Vision-Based Approach

Yörük et. al. [17] developed a method for real-time hand based person recognition which works with unconstrained hand poses and no restrictions on hand accessories, i.e. no special platform for hand alignment is required; hands are registered to a fixed pose by rotating and translating them. In their experiment, they used a flatbed scanner with a resolution of 45dpi and resampled the retrieved images to 20dpi after pose normalization. Today's computer vision cameras applied in tabletop systems are capable of delivering similar resolutions; for larger tables several cameras can be combined.

The best performing algorithm tested in this study yields correct classification rates of promising 97%, that is for a population of sizes within hundreds. It is important to note that users have to keep their fingers apart, though. As there is only a limited number of people interacting at the same time around a tabletop, our hypothesis is that fewer features are sufficient to distinguish between these users. Features can be collected by tracking the hand while the user is interacting above the surface, using different hand shapes over time. The most commonly used hand shape is pointing with the index finger (70%) followed by a spread hand (20%) [5] which can potentially be used to derive proportional information about fingers and hands as distinguishable features. Alternatively, the user could be asked to put his or her hand with the fingers spread on the surface in order to perform an identification.

In order to test the concepts discussed before, we built a DI based rear-projected tabletop system with an active surface of about 90cm $\times$ 56cm. Touches are picked up by a camera mounted below the surface. As we use a diffusing surface to make touches visible and project onto the screen, hand shapes are not clearly visible anymore. Therefore, we installed an additional overhead camera pointing down at the table's surface. Both cameras use infrared bandpass filters to eliminate disruptions from environmental light sources. Moreover, this approach facilitates the extraction of hand contours since the hands cast clear shadows in front of the rear-illuminated surface. In future systems, a switchable diffuser [7] may be used to realize a more compact system with only one camera.

FUTURE WORK

Applying the proposed concept to the system we built, we aim at answering the following questions: Are the features that can be extracted out of the hand shapes during non-constrained interaction on the surface distinguishable enough to perform a clustering in order to differentiate between the hands of users? How do the applied computer vision methods deal appropriately with obstacles, such as overlapping hands? Moreover, we are interested in the user acceptance of an explicit identification gesture – such as putting the flat hand on table. And last but not least: To what extent are users concerned about privacy issues, considering that biometric mechanisms are applied?

ACKNOWLEDGEMENTS

This research was conducted with support of the European Commission, as part of the Research Network of Excellence "InterMedia".

REFERENCES

1. E.A. Bier and S. Freeman. MMM: a user interface architecture for shared editors on a single screen. In *UIST '91*, pages 79–86, 1991.
2. J. Carreira and P. Peixoto. Retrieving and exploiting hand's orientation in tabletop interaction. In *Multimedia and Expo, 2007*, pages 1003–1006, 2007.
3. P. Dietz and D. Leigh. Diamondtouch: a multi-user touch technology. In *UIST '01*, pages 219–226, 2001.
4. K. C. Dohse, T. Dohse, J. D. Still, and D. J. Parkhurst. Enhancing multi-user interaction with multi-touch tabletop displays using hand tracking. In *Advances in Computer-Human Interaction, 2008*, pages 297–302, 2008.
5. J. Epps, S. Lichman, and M. Wu. A study of hand shape use in tabletop gesture interaction. In *CHI '06*, pages 748–753, 2006.
6. Circle Twelve Inc. DiamondTouch flyer, 2008.
7. S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display: A surface technology with an electronically switchable diffuser. In *UIST '08*. ACM Press, 2008.
8. B.A. Myers, H. Stiel, and R. Gargiulo. Collaboration using multiple pdas connected to a pc. In *CSCW '98*, pages 285–294, 1998.
9. K. Ryall, A. Esenther, K. Everitt, C. Forlines, M. R. Morris, C. Shen, S. Shipman, and F. Vernier. iDwidgets: parameterizing widgets by user identity. pages 1124–1128. 2005.
10. K. Ryall, A. Esenther, C. Forlines, C. Shen, S. Shipman, M.R. Morris, K. Everitt, and F.D. Vernier. Identity-differentiating widgets for multiuser interactive surfaces. *IEEE Comput. Graph. Appl.*, 26(5):56–64, 2006.
11. S.S. Scott and S. Carpendale. Guest editors' introduction: Interacting with digital tabletops. *Computer Graphics and Applications, IEEE*, 26(5):24–27, 2006.
12. C. Shen, K. Everitt, and K. Ryall. UbiTable: impromptu face-to-face collaboration on horizontal interactive surfaces. pages 281–288. 2003.
13. J. Stewart, B.B. Bederson, and A. Druin. Single display groupware: a model for co-present collaboration. In *CHI '99*, pages 286–293, 1999.
14. E. Tse and S. Greenberg. Rapidly prototyping single display groupware through the SDGToolkit. In *AUIC '04*, pages 101–110, 2004.
15. E. Tse, C. Shen, S. Greenberg, and C. Forlines. Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. In *AVI '06*, pages 336–343, 2006.
16. M. Wu and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *UIST '03*, pages 193–202, 2003.
17. E. Yoruk, E. Konukoglu, B. Sankur, and J. Darbon. Shape-based hand recognition. *Image Processing, IEEE Transactions on*, 15(7):1803–1815, 2006.